Public Discourses on Multilingualism in the UK: Triangulating a corpus study with a sociolinguistic attitude survey

INTRODUCTION

Multilingualism is a contentious issue. Although linguistic diversity is generally celebrated (e.g. Milani, Davies, & Turner 2011), it is also considered problematic (Blackledge 2004; Horner 2011). The now considerable body of research concerned with representations of multilingualism has shown that the mainstream media are the key agents in disseminating contentious and ideologised representations of multilingualism (Johnson & Ensslin 2007; Androutsopoulos 2010; Horner 2011; Jaffe 2011; Kelly-Holmes & Milani 2011). Despite the breadth and depth of this research we believe that there remain three challenges warranting research attention.

Firstly, most research in this area focuses on representations of a particular language and proceeds inductively to draw general conclusions about multilingualism as a societal phenomenon (e.g. Androutsopoulos 2010; Horner 2011; Jaffe 2011; Kelly-Holmes 2011). While rich in findings, each study illustrates a specific case, revealing views about multilingualism that might not be applicable to other contexts. We believe therefore that there is an advantage in complementing this research by using a deductive approach which starts with the concept of multilingualism. In other words, we propose to examine how the term *multilingualism* is discursively constructed in public domains to identify which languages and practices are foregrounded or backgrounded. This is especially relevant in superdiverse contexts such as many urban centres of the UK where multilingualism is not linked to any particular combination of languages.

Secondly, many studies concerned with mainstream media representations (not just of language) assume a causal relationship between media and audiences, where the media are

1

described as representing or influencing public attitudes. While mainstream media representations are public in the sense that they are in the public domain, they might not necessarily reflect what the public thinks. Rather, the mainstream media as institutions of power are linked to political and cultural elites and tend to reflect views that are representative of those in power (Fairclough 1989). Investigating mainstream media representations of multilingualism and comparing them with what lay people think about these representations could potentially help us unpack the relationship between the mainstream media and the public, thereby revealing which representations reflect public understanding and which are contested. We consider such an approach imperative today, when following political shock events such as Brexit or the US election, the mainstream media, political, and cultural elites have all been criticised for being disconnected from the views of 'people on the street' (e.g. Blagden 2017).

Finally, most studies concerned with media representations of multilingualism are based on small amounts of media texts collected at a particular point in time offering rather snap-shot views of how multilingualism is conceptualized. Fairclough (1989: 54) reminds us that 'the effects of media power are cumulative, working through the repetition of particular ways of handling causality and agency, particular ways of positioning the reader, and so forth'. Understanding which messages about multilingualism are repeated across mainstream media necessarily requires large data sets collected over a longer period of time. Here, corpus tools and methods are useful as they can help reveal patterns of repeated discourses in larger amounts of data (Baker, Gabrielatos, & McEnery 2013; Vessey 2016).

We propose to respond to these challenges by examining representations surrounding the term *multilingualism* in a large corpus of articles published in British national newspapers from 1990 to 2014. Our understanding of multilingualism is embedded within poststructuralist research; following Heller (1999) and Martin-Jones, Blackledge and Creese (2012), we see multilingualism as a form of symbolic capital (Bourdieu 1991) and a linguistic resource that is

unequally distributed in society and whose meanings and values are discursively constructed depending on social, political and historical conditions. We are therefore specifically interested in *what is frequently said* about multilingualism and *how is said*, as well as how these discourses may have changed over time. To investigate this, we use the approach of Corpus-Assisted Discourse Study (CADS) (Partington, Duguid, & Taylor 2013) which is increasingly adopted for the study of media representations. Discourses identified using CADS were subsequently included in a sociolinguistic attitude survey distributed to nearly 200 participants. Using this triangulation of methods, our study offers a much more nuanced understanding of the ways in which multilingualism is 'thought of' in the public domain and how views have changed over time. The study also offers important pointers to teachers, campaigners and organisations who strive to promote multilingualism in society.

RESEARCH ON PUBLIC DISCOURSES ON MULTILINGUALISM

Theoretical concepts

   Before we proceed with an overview of the research literature on the topic, we would like to briefly discuss two theoretical notions that have influenced much research about representations of multilingualism and that are central to this study. These are *metalanguage* and *mediatisation*.

   Although often understood simply as 'language about language' (Jaworski, Coupland, & Galasiński 2004), metalanguage is a multifaceted construct with several components. Preston (2004) makes a distinction between *Metalanguage 1*, *Metalanguage 2* and *Metalanguage 3*. Whereas Metalanguage 1 refers to explicit comments on language use, Metalanguage 2 involves references to the talk itself (e.g. *in other words*). Metalanguage 3 includes folk beliefs and attitudes towards language use that, as Preston (2004: 87) observes, are widely shared presuppositions and inextricably linked with Metalanguage 1. Such beliefs are not just a matter

of personal opinion, but manifestations of deeply-rooted language ideologies. They are 'common-sense' views about what counts as legitimate language use and which establish links between language feature(s) and non-linguistic factors such as, for example, speaker's sex and educational background (Irvine & Gal 2000). When frequently repeated, such links can end up being collective 'truths' (Blommaert 2005) with real and often degrading effects on speakers. This happens through various semiotic strategies, such as iconisation, erasure, fractal recursivity (projections of differences) (Irvine & Gal 2000), commodification (Heller 2003), festishisation (Kelly-Holmes 2011) or othering (Androutsopoulos 2010).

By investigating what is said about multilingualism and how it is said in public domains (that is, the public metalanguage of multilingualism), we aim to demonstrate the extent to which (language) ideologies influence how multilingualism is conceptualised and to show the mechanisms through which such ideologies become manifest. The mainstream media present a useful source of data here because they are key agents in ideologizing multilingualism (Androutsopoulos 2010).

The second concept, that of mediatisation, offers a useful perspective within which to conceptualise the link between the media and the public (Androutsopoulos 2014). In relation to language in the media, mediatisation is mostly understood as 'all the representational choices involved in the production and editing of text, image, and talk in the creation of media products about language' (Jaffe 2011: 98). Agha (2011: 163) emphasises the relationship between media messages and the various forms of uptake or recontextualisation of the messages by audiences. He stresses that audiences are not empty receivers eager to accept whatever the media tell them. Rather, media provide mass inputs that are further recontextualised. The notion of mediatisation has been shown to be productive when exploring the relationship between media and language change (Androutsopoulos 2014). However, little is known about the public uptake of metalinguistic discourses about multilingualism (see, however, Jaffe 2007, Kelly-Holmes and

Milani et al 2011). In this study we 'expose' a large number of monolingual and multilingual respondents to dominant media messages about multilingualism and analyse their responses to such views, thereby revealing the sometimes contradictory and unexpected ways in which mediatised representations of multilingualism are recontextualised by 'people on the street'.

Multilingualism in the media

Research into media representation of multilingual practices sprang from the sociolinguistic interest in language in the media which began in the late 1990s, inspired by studies on language ideology (Johnson & Ensslin 2007). Focusing on multilingual Luxembourg, Horner (2011) examines media discourses about the country's educational system, demonstrating how the trilingual ideal, centred on literacy in German and French and spoken Luxemburgish, is projected as 'proper' multilingualism and representative of 'real' Luxembourgers. Conversely, multilingualism based on different language combinations, for example those used by migrant children, is seen as problematic. Jaffe's research in Corsica (2007) shows how local media construct linguistic varieties revealing and purposefully reconstruct a bilingual community as monolingual and homogenised. Turning to Germany, Androutsopoulos (2010) investigates media representations of *Kanak Sprak* - a xenophobic term used to describe a youth ethnolect of multilingual urban speech communities. He shows how the ideology of standard German is constantly evoked to iconise *Kanaksprak* as a kind of 'bad' German indexing migrant youth as socially problematic and not-integrated. In a similar vein, Kerswill's (2014) work on the representations of Jafaican (a lay term for Multicultural London English, the multiethnolect spoken in London) in the British mainstream press demonstrates how this term was coined and repeatedly used by the British press as an index of foreignness, bad behaviour and a threat to 'Englishness' and social cohesion.

Attitudes towards multilingualism in the UK are of interest to Milani et al (2011) who investigate the website of the BBC Voices project, set up with a view to collect examples of regional English accents and dialects. They show how the BBC representations of multilingual Britain are essentially centred on the 'monolingual voice' of Standard British English, while the diversity of multilingual speakers is largely erased. Such a representation was challenged in online postings commenting on the use of Welsh. Some posters from Wales saw Welsh as a marker of their national identity and thus resisted the BBC's view of the UK as a nation united in its diversity. This study is one of the first examples showing how the views of lay people can challenge discourses promoted by media elites.

Kelly-Holmes (2011) also explores views of the general public, focusing on discourses about Irish in an online discussion forum. While some of the discourses she identifies confirm the stereotype that Irish is 'inferior', she also noticed the emergence of a new type of discourse centred on the notion 'Irish is sexy'. In her view, this newer discourse promotes bilingualism-as-an-added-value, counter-balancing the pervasive monolingual ideal, specifically the view that 'English only' can be of benefit to Ireland. However, this positive discourse is a rather fringe phenomenon. Moreover, most of the online posts seem to endorse the idea of parallel monolingualisms (Heller 1999), while everyday bilingual practices so typical of the sociolinguistic reality of Ireland are largely erased.

Discourses commodifying bilingualism as an added value are not new; they have been observed in other bilingual contexts (Duchêne & Heller 2011). After all, knowledge of languages that have a high status have long been marketed as an asset. However, commodification seems to have accelerated in our globalised economy, which places immense value on symbolic capital and distinctions. In this process even some lesser spoken languages have shifted their status from being a marker of ethnonational identity to being a marketable commodity promising economic returns (Duchêne & Heller 2011; Kelly-Holmes & Pietikäinen

2014). It is striking that although indigenous minority languages have experienced this kind of positive revalorisation, languages spoken by migrants are still largely 'unaffected' by this shift. Blommaert et al. (2009: 205) remind us that: 'National minorities […] have linguistic rights, while immigrants do not and are thus not included in the multilingual spectre of the state'.

The studies surveyed above point to several distinctive patterns that pervade mainstream media representations of linguistic diversity. Firstly, in contexts where multilingual practices are reported, the dominant frame of reference seems to be that of monolingualism based on a standard national language (Androutsopoulos 2010; Milani et al 2011), while multilingualism is mostly conceptualised as parallel monolingualisms (Jaffe 2007; Kelly-Holmes 2011). This is reinforced through the reduction of complex multilingual practices to a few essentialising and often pejorative images that degrade multilingual speakers (Kerswill 2014). Multilingual practices such as code-switching, in which multilingual speakers engage on a regular basis, are hardly ever mentioned (Kelly-Holmes 2011). This is a manifestation of the inequality which persists in the linguistic market (Bourdieu 1991): the ability to speak high status national languages, including some minority languages, is commodified as 'proper' multilingualism, whereas speaking 'low status' languages is considered problematic.

However, while studies such as these deal with dominant media discourses about many facets of multilingualism, little is known about the extent to which these discourses are taken up by the general public. Research into language attitudes can offer some useful pointers.

Public attitudes towards multilingualism

Although the study of language attitudes towards multilingualism is relatively new (Lasagabaster & Huguet 2007; Baker 2008; Caruana & Lasagabaster 2011; Dewaele & Li 2014; Stavans & Hoffman 2015), some research has provided invaluable insights into the ways in which language attitudes are inextricably linked with mechanisms of inclusion and exclusion

as well as patterns of cultural assimilation or resistance. This section discusses a selection of indicative studies in this area.

Within the area of language attitudes, attitudes towards foreign accents have received a great deal of attention. For example, Garrett (2010) shows that speakers with foreign accents are likely be judged as less educated, less competent and less trustworthy compared to those with native accents. Two factors seem to influence attitudes towards accented speech: the status of the language which 'causes' the accent (Giles 1970) and personality traits of the judges (Dewaele & McCloskey 2015). Dewaele and McCloskey (2015) have shown that multilinguals who are extravert, emotionally stable and tolerant of ambiguity can be significantly less bothered by foreign-accented speech.

Attitudes towards code-switching have also been widely explored. Code-switching tends to be perceived negatively; it is associated with laziness, bad manners, language inability or poor cognitive control (Garrett 2010). Gardner-Chloros, McEntee-Atalianis and Finnis (2005) found that age, occupation and level of education can affect attitudes towards code-switching too. They show that negative attitudes were more likely to be expressed by older, educated participants with higher status jobs. Dewaele and Li's (2014) work also reveals that socio-biographical factors can affect attitudes towards code-switching. For instance, positive attitudes towards code-switching are expressed by multilingual individuals who grow up in a multilingual family and live and work in ethnically diverse environments. Attitudes may also depend on the context in which multilingual practices occur. For example, Balam and de Prada Pérez (2017) report positive attitudes towards code-switching held by teachers of Spanish who viewed this practice as an important pedagogical tool in the classroom.

Finally, some scholars have investigated attitudes towards majority vs. minority languages in multilingual contexts (e.g. Hoare 2000 on Breton; Pritchard 2004 on Irish; Novak-Lukanovič & Limon 2014 on Italian and Hungarian in Slovenia). As immigration is closely related to the

emergence of multilingual communities, some scholars also compare the attitudes of 'local' vs. immigrant people. For instance, Ibarraran, Lasagabaster and Sierra (2008) show that both local and immigrant students in the Basque Country hold more negative attitudes towards Basque compared to Spanish and English, which they viewed as more useful. Within the context of the UK, Hilmarsson-Dunn and Mitchell (2011) investigated attitudes of monolingual and multilingual (immigrant) students and teachers towards multilingualism. They showed that attitudes towards English are positive, as English is considered important for becoming assimilated in the community. At the same time, however, students endorsed multilingualism and viewed it as an opportunity to develop better careers. Teachers accepted the use of other languages in school but insisted that the main language should be English.

Evaluating methodological approaches used to investigate language attitudes, Caruana and Lasagabaster (2011) highlight the rather one-sided methodology that tends to be used and encourage researchers to use a more holistic approach, including a range of socio-biographical variables (see, for example, Gardner-Chloros et al 2005). It is striking that research in this area is normally based on monolingually-biased questionnaires focusing on the binary opposition between a majority and minority language. While this may be relevant in strictly bilingual environments, arguably this kind of research has fewer implications for superdiverse contexts. Also, most studies are interested in the impact of multilingualism on education and are therefore concerned with one particular professional group – teachers – and students, with other groups rarely included. In addition, very little is known about how the dominant media discourses about multilingualism might influence public opinions. By combining media perspectives with public attitudes, our research adopts a holistic approach and provides a more comprehensive understanding of public views about multilingualism.

RESEARCH METHODOLOGY

This study follows a two-tier approach. To identify the dominant and frequently repeated discourses about multilingualism in the British press we use the CADS approach (Partington et al 2013) and interrogate a large corpus of articles published in the major British national newspapers. Results obtained from this part of the study were then incorporated in the design of an attitude survey which was distributed to almost 200 participants living in a large and ethnically diverse town in the south of England. We first offer a detailed description of the corpus methodology and then outline the design and administration of the survey.

Corpus Methodology and Data

CADS is based on a combination of corpus tools and methods with qualitative procedures commonly adopted in discourse analysis. The benefit of Corpus Linguistics lies in its capacity to reveal, through keywords and collocations, repetitively occurring lexico-grammatical patterns which, in turn, can point to salient representations and majority ways of viewing the studied phenomena (Baker et al 2013). CADS is not confined to any specific type of discourse analysis, thereby allowing the researcher to engage with the corpus in many different ways (Partington et al 2013). We follow this practice in that we employ quantitative tools to study frequent representations surrounding the term *multilingualism*. Selected keywords and collocations are subsequently examined qualitatively by studying concordance lines and text extracts.

This part of the study interrogates a large corpus consisting of press articles published in 12 major British national newspapers between 1 January 1990 and 1 May 2014, named *Multilingualism in Public Discourse* (MinD) (see Appendix 1). Since we are interested in discourses widely disseminated across the UK we decided to include national newspapers with wide circulation, both broadsheets and tabloids. Articles with the search terms *multilingual!*

and *bilingual!* were subsequently downloaded from Nexis UK. The punctuation mark '!'
replaced any set of characters and was used to retrieve articles in which other word forms
containing *multilingual* and *bilingual* occurred, such as *bilinguals* and *multilingualism*. We
included *bilingual!* in our searches because this term is commonly used to describe multilingual
practices. The terms *multilingual* and *multilingualism* when used in this study incorporate
bilingualism. To ensure that multilingualism was topical and not mentioned only in passing,
only articles in which these terms occurred at least 3 times were included. Each article
downloaded from Nexis UK comes with metadata including the publisher, the date and place
of publication. This information could potentially inflate the results and was removed using
regular expressions and the editor Notepad++[i]. In order to investigate changes over time the
corpus was divided into 3 subcorpora, each representing a different decade. Table 1 summarises
the number of articles and words in each subcorpus.

Table 1 about here (see Tables at the end of the document)

To identify dominant discourses about multilingualism in each decade we performed a
keyword analysis using Sketch Engine. Keywords are generally considered good indicators of
the 'aboutness' of texts and the salient themes in a given data set. In corpus linguistics terms,
'keyword' is a word which occurs unusually frequently in a given corpus as compared to
another reference corpus (Scott 2010). Corpus-based retrieval of keywords often utilises the
British National Corpus (BNC) as a reference corpus because it is regarded as a representative
compilation of British English. We too used the BNC as our reference corpus. The outputs are
normally presented in the order of keyness established using either log likelihood as the measure
of statistical significance (Scott 2010) or a ratio of normalised frequencies (Kilgarriff 2005).
Because our data sets are of unequal sizes, we used the method suggested by Kilgarriff (2005),

as it does not depend on significance testing, which in turn relies on the sample size. Keywords with the highest keyness scores are normally seen as distinctive of a given data set.

To capture the main themes the first 100 keywords were subsequently grouped into semantic categories – a procedure adopted in previous research on media representations (e.g. Gabrielatos & Baker 2008; Baker, Gabrielatos & McEnery 2013). Subsequently, we investigated a selection of keywords in context by examining their collocations. The selection included keywords that are present in the three corpora and hence are consistently associated with multilingualism, such as *English* and *bilingual,* as well as one new keyword, *Alzheimer,* which was identified as representative of a new discourse. Collocations were retrieved using a -5 to +5 span and log dice (LD) as the measure of statistical association. We selected LD as a metric of association because in contrast to other widely used metric, such as Mutual Information or t-test, LD is a ratio with a maximum value (theoretically 14) and does not depend on the total size of the corpus (Rychlý 2008). This allows us to have a consistent comparison measure across our data sets.

Design and Administration of the Attitude Survey

The second part of the study was based on a holistic survey, which was distributed to residents of Reading. Reading is a large town in England with a population of 155,698 and the second town with the highest population of 'non-white' ethnic groups in the South East of England (Office for National Statistics 2011). The top 3 countries of birth outside Great Britain are India, Poland and Pakistan, with a large proportion of residents coming from African countries as well as elsewhere in the EU. Reading therefore was considered to be an ideal superdiverse town in which to conduct the survey because it is representative of the demographic changes that have occurred in the UK post the EU expansion in 2004.

The questionnaire was divided into three parts. The first part gathered demographic information about the participants' sex, age, ethnic background, duration of residency in the UK and their native language (i.e. the first language they acquired). The question *'Do you consider yourself multilingual (i.e. know/speak more than one language)?'* was used to categorise monolingual and multilingual participants in our sample and the follow-up question *'If yes, what other languages do you know/speak?'* was used to obtain information about the number of languages used.

The second part focused on participants' attitudes towards living in a multilingual speech community, prompted by the question: *'Overall, would you say that living in a town like Reading where lots of languages are spoken is a positive or negative thing?'*. Participants could select from three options: *positive*, *negative* or *both positive and negative*. This section was also an excellent source of qualitative data, as participants were invited to give reasons for their opinion.

The third part explored participants' attitudes towards the dominant media discourses about multilingualism found in the corpus analysis. It consisted of a series of statements including the following discourses or themes: 'formal education', 'elite multilingualism', 'employment', 'medical' and 'multilinguals as non-native speakers of English' (see Table 14 below). Participants were asked to rate their agreement or disagreement with each statement using a 5-point Likert scale. The last question presented the participants with five profiles of multilingual speakers focusing on their competency in his or her languages, qualification(s) in another language and employment. The languages used in the profiles were the 'prestigious' languages consistently associated with multilingualism in the press, that is, English, French and German, as well as languages that have been featured more often in recent years, such as Polish and Urdu. Participants were asked to indicate whether they thought the person in each scenario was

multilingual or not by choosing *yes*, *no* or, if they were unsure, *perhaps* (see the section *Who is multilingual?*).

Data collection took place in Reading city centre in May-June 2014. This area was chosen as it is one of the most ethnically diverse parts of Reading (Office for National Statistics 2011). Using the convenience sampling technique, the fieldworker approached participants and asked them to complete the questionnaire. In total, 194 participants took part in this survey (see Table 2). All data was processed using SPSS. In order to estimate whether there is a relationship between participants' attitudes and the socio-biographical variables included in the survey we used the Chi-square test.

We acknowledge that because our sample is not socially stratified some categories in our dataset are over- or underrepresented; for instance, there are more females in our study than males. This was because females were more willing to speak to the fieldworker and complete the questionnaire. However, following Wilson and Dewaele (2010), we did not consider this to be a major problem as for multilingualism research it is more important to have participants who are keen to engage with the research than to have a perfectly matched sample. Also, the advantages of using a convenience sample might outweigh the disadvantages. For example, in comparison to previous attitudinal studies which targeted mostly teachers and students, our sample was more diverse in terms of age and ethnic background. Secondly, we were successful in collecting data from almost an equal number of monolingual and multilingual speakers. The latter group was very diverse, representing 24 different languages (Arabic, Bulgarian, Chinese, Creole, Danish, Dutch, French, German, Greek, Marathi, Nepalese, Persian, Portuguese, Polish, Punjabi, Romanian, Russian, Sinhalese, Spanish, Tagalog, Tamil, Thai, Turkish and Urdu) and thus in many ways representative of superdiverse contexts in Britain.

Table 2 about here

RESULTS AND ANALYSIS

This section summarises the major results of the study; we first discuss findings obtained from the corpus media analysis and then summarise the results of the attitude survey.

Media discourses about multilingualism

Table 3 shows the first 15 keywords retrieved from each subcorpus. As can be seen, most of the keywords are the same, highlighting constant themes in the three decades. The salience of *English* is striking but perhaps not surprising. The second constant language is *French*, which is due to its status as the first foreign language in British education. Overall, there seems to be a heavy focus on the domain of education, as most of the keywords point to aspects of schooling. Lesser spoken languages such as Welsh and Gaelic are keywords in MinD1 and MinD2, but disappear from MinD3. This might suggest that regional languages are nowadays given less attention when multilingualism is discussed in the media. The appearance of *Alzheimer* in MinD3 is interesting, suggesting a slight degree of 'medicalisation' of bilingualism, an issue to which we return below.

Table 3 about here

Following Gabrielatos and Baker (2008), we grouped the first 100 keywords into semantic domains to see which themes dominated in each decade. Keywords are isolated lexical items that might have different meanings depending on the context. To ensure that the keywords were grouped into the appropriate semantic category, we checked the use of the keywords in the articles to see their actual meanings in context. Tables 4, 5 and 6 present some examples of keywords identified in each semantic domain. Because of space restrictions, the list is not exhaustive.

Table 4 about here

Table 5 about here

Table 6 about here

The tables point to some constant themes, but also some themes that shift across time. The most salient constant theme is that of formal education, as indicated by the many keywords from the domain of schooling. Thus, multilingualism seems to be viewed predominantly as something to be developed as part of formal school education; it is not necessarily associated with growing up in a bilingual community. This is further reinforced by the prominence of *teachers*, *pupils* and *students*; these are the constant social actors that occur across the three subcorpora. Against this background, it is not surprising to see that multilingualism is also consistently associated with prestigious foreign languages that are school or university subjects, including French, German and Spanish. Another constant pattern is the occurrence of keywords referring to large metropolises such as London or Paris. The only exception is Peterborough, a comparatively smaller town located in Cambridgeshire, in the East of England. We will return to this case below.

Alongside constant keywords there are a few items that occur only during one or two decades, thus suggesting some thematic shifts. This concerns keywords from the medical domain, specifically references to neurodegenerative diseases of the brain (*Alzheimer*, *dementia*), which occur unusually frequently in MinD3 but are absent from MinD1 and MinD2. Shifts can also be observed in the domains of employment and languages. Whereas in MinD1 we find several keywords pointing to employment opportunities occur, they play lesser role in MinD2 and MinD3. Similarly, two important languages of the UK, Welsh and Gaelic, are identified amongst the first 100 keywords in MinD1 and MinD2 but not in MinD3. Instead, we

find here new languages such as Polish and Urdu. This could suggest that nowadays multilingualism is associated less with minority languages than was the case in the 1990s. The sudden keyness of *Polish* reflects recent demographic changes in the UK: according to the 2011 census Polish is now the third most frequently spoken language in the UK. Urdu too is one of the most spoken languages, reaching the fifth place after Polish and Punjabi, but, interestingly, Punjabi was not identified as a keyword. Some topical shifts also occurred in the domain of social actors. From 2000 onwards we can notice the occurrence of *immigrants* as well as the related term *immigration* amongst the first 100 keywords.

While keywords are useful pointers to general topics, they are rather 'blunt instruments' (Gabrielatos & Baker 2008: 28) that on their own tell us little about discursive meanings in context. Thus, a keyword analysis needs to be expanded by examining the actual use of keywords in texts via concordance lines. Each keyword presented in the tables above is a good candidate for an in-depth analysis; but as space is limited, we selected just 3 keywords. Two were judged representative of constant trends (*bilingual*, *English*) and one of shifting patterns (*Alzheimer*). Keeping in mind the aim of the survey, we intended to gather data about the most recent discourses surrounding multilingualism; hence, only texts sampled in the MinD3 corpus were examined.

We begin with collocations of *bilingual*. As Table 7 shows, *bilingual* collocates strongly with items pointing to education, especially the primary sector. We assume that this relates to the recent introduction of a foreign language component in British primary schools. The modifier *better* refers mostly to the better performance of bilinguals on certain mental tests

Table 7 about here

when diagnosing dementia. *Bilingual* is also strongly associated with *English*. Lower down on the list we also find *French* and *Spanish*. These patterns confirm the tendency to associate multilingualism with formal education and prestigious languages. Languages of large immigrant populations, such as Polish or Urdu, are not the strongest collocates of *bilingual*. Table 8 shows the 10 strongest associations with *English*. As can be seen, *English* is mostly associated with the ability to speak the language (*speak*, *use*). It also collocates with another

Table 8 about here

prestigious language, *French*. Like *bilingual*, languages occurring in the vicinity of *English* are high status languages; community languages do not collocate with the term. Another striking result is the collocation *their* and *not*. Studying the concordance line of the collocation pair *English* and *their* reveals that in most cases (47 out of 66), *their* was followed by a quantifier including *first*, *second* and *main*, forming clusters such as *English as their first/second/main language*. Hence, English is predominantly framed as a language spoken by speakers who have other first languages (L1); but, interestingly, in most cases the other L1s are not mentioned. This becomes evident when looking at the concordance lines of the collocation pair *first* and *English*. The collocation occurs 50 times in the corpus, of which 24 instances include *not* (13 times), *none* (6 times) or *don't'/doesn't'/did not* (5 times). Figure 1 shows illustrative examples. In most instances, *speak* is negated and a specific group of people, that is, children who do not speak English as their first language, are foregrounded.

Figure 1 about here

Another striking result is the frequent use of quantification (e.g. *not a single*, *80.8%*) when referring to *children*. Quantification is often used as a rhetorical device to construct demographic groups by emphasising certain criteria and obscuring others (Jones 2013). In this case the numbers are deployed to amplify the message that there are large numbers of children in British schools who do not speak English as their L1, while their multilingual abilities are erased. Thus, multilingualism is seen here as a problem rather than an opportunity. Yet this discourse is not shared by all newspapers. It seems to be typical of the middle-range tabloid the *Daily Mail,* as all the negative examples come from this source. Extract 1 below featuring a school in Peterborough is representative. Even though a positive view about multilingualism is expressed, this is immediately qualified by pointing to financial challenges and 'problems' with literacy and numeracy, for which little evidence is, in fact, provided.

Extract 1:

"A positive view of the bilingual child is essential,' says Mrs Parker, 54. They are an asset, not a liability. […] But teaching 445 children who speak 23 languages is a job full of expensive and time-consuming challenges. It also raises the question where are all the white English-speaking children? In truth, most white British families have moved out of its catchment area. Many of Gladstone's pupils arrive speaking no English, or a muddled smattering, at best. More taxpayers' money is spent to provide one-on-one help for any children having difficulty with reading and numeracy." *Daily Mail*

The same situation is described in *The Guardian,* which is considered a politically left-leaning newspaper. However, as Extract 2 shows, the school is portrayed differently here: the good results and a competent use of English by bi- or multilingual children are emphasised.

Extract 2:

Gladstone Primary in Peterborough doesn't have one pupil who speaks English as a first language. But, despite the challenges, it has just received a glowing Ofsted report […] the children huddled around tables wouldn't alarm the most bigoted of columnists: they speak perfect English.

Although *The Guardian* offsets the alarmist stance expressed in the *Daily Mail*, it is the *Daily Mail* which is one of the bestselling newspapers in the UK, with 5 million readers. The readership of *The Guardian* is more modest by comparison, with around 1 million readers. Thus, the views expressed in the *Daily Mail* that link societal multilingualism with immigration and problems are likely to be more widely disseminated.

Finally, we consider the collocations of *Alzheimer*, which seem to represent a new type of discourse about multilingualism. As Table 9 shows, *Alzheimer* is strongly associated with diagnosis of this disease and one may reasonably ask what this has to do with multilingualism. Studying the concordance lines of the 10 collocates reveals that nearly all collocates occur in

Table 9 about here

relation to research conducted by the cognitive psychologist Ellen Bialystok and her collaborators on the effects of bilingualism on the delayed onset of Alzheimer's. This research provides some evidence for a four-year delay in symptom onset in bilingual speakers and suggests that bilingualism might contribute to the cognitive reserve that compensates for the effects of accumulated neuropathology. In most instances the reporting presented the facts accurately. Yet the headlines were at times less exact, relying on the use of war metaphors such as *keep at bay*, *combat* and *fight off* (see Extract 3). The use of war and military metaphors in discussing medical conditions is not unusual (e.g. Demmen et al 2015). In the cases below the use of such metaphors serves to frame bilingualism as a 'weapon' against Alzheimer's, creating the impression that bilingualism is a cure for the illness, which it is not.

Extract 3:

a) A second language can keep Alzheimer's at bay (*The Times*)

b) Speaking a second language increases "brain power" (*Daily Telegraph*)

c) Being bilingual can help stave off dementia and protect against Alzheimer's disease (*The Sun*)

d) Languages combat dementia (*The Observer*)

The corpus-assisted analysis above has shown some dominant patterns in the representations of bi- and multilingualism in the British press. Generally, we can identify two salient discourses: one which sees multilingualism as an opportunity and value, and one which frames it as a problem. Multilingualism-as-an-opportunity discourse is centred on the notion of educational bilingualism, strongly associated with formal schooling and prestigious languages such as French and Spanish. This reinforces the notion of an elite bilingualism which values prestigious languages, formal education and qualification, and which devalues or excludes languages spoken by immigrants. The absence of community languages as collocates of *bilingual* is indicative of the dominance of elite bilingualism. Multilingualism-as-a-value discourse is associated with employment opportunities (though less so in recent years) and with positive impacts on the delayed onset of Alzheimer's. However, the positive effect is, at times, exaggerated, creating the myth that bilingualism could 'cure' the disease. The bilingualism-as-a-problem discourse is mostly based on the notion of English as 'not the first language' of large groups of immigrant children, and features predominantly in the right-leaning press. The multilingual abilities of immigrant children are hardly ever mentioned. Instead, multilingual children are framed as immigrants and linked with burdens on public services and issues with literacy.

PUBLIC ATTITUDES TOWARDS MULTILINGUALISM

Attitudes towards living in a multilingual speech community

Responses to the question '*Overall, would you say that living in a town like Reading where lots of languages are spoken is a positive or negative thing?*' indicate that the majority of the participants (65%) held positive attitudes, while 25% expressed both positive and negative views, and 10% expressed exclusively negative opinions (see Figure 2).

Figure 2 about here

The open question that followed prompted participants to elaborate on their choice and provided further insights into the motives underlying the reported attitudes. Some indicative answers are presented in examples (1) – (9) below.

Positive opinions:

(1) **Participant 68:** 'It shows that people are more aware of diversity and able to perceive other people's traditions and cultures (i.e. bring mutual respect and understanding between nations and bring peace and healthy co-existence)'

(2) **Participant 90:** 'Many languages enriches the cultural landscape making us a more tolerant society'

(3) **Participant 171:** 'It denotes an open minded society which respects different people. More importantly, it is more creative at all levels (economy, education)'

Negative opinions:

(4) **Participant 47:** 'Only people with native English must work in call centres, nurseries, any public service where it's necessary to speak to customers. Sometimes when I call NHS people with a strong Indian accent reply. I hardly can understand them.'

(5) **Participant 54**: 'They steal jobs and are annoying.'

(6) Participant 57: 'They need to learn the language of the country they are in instead of making us feel like we are abroad.'

Mixed opinions

(7) Participant 2: 'I'm ambivalent! A good mix of cultures and languages is a positive thing but sometimes I walk down Friar Street and don't hear any English + feel rather indignant!'

(8) Participant 21: 'I believe diversity is a good thing and multi-culturalism can strengthen a community. The flip side to this is that if there is a distinct language barrier between communities then it can cause those communities to become isolated'

(9) Participant 132: 'I think coming into contact with different languages and different cultures on a daily basis can only be a good thing for people. However, I do believe there can be difficulties when people come to live here from other countries and do not try to learn English.'

The majority of those who held positive attitudes towards living in a multilingual speech community emphasised cultural diversity, opportunities for learning new languages and cultures by mixing with speakers of other languages. Many also expressed the opinion that multilingualism makes people more open-minded and tolerant. On the other hand, participants who expressed negative opinions tended to link multilingualism with difficulties in communication, immigration and loss of jobs and, overall, with inequality and conflict. Although the question prompted the participants to comment only on their views about living in a town where many languages are spoken, respondents who expressed only negative views immediately linked multilingualism with immigration and immigrants, the latter often referred to using the pronoun *they*. This usage is a good example of 'othering', which emphasises the

distinction between 'they' (i.e. immigrants, speakers of languages other than English) and 'us' (i.e. the local, British English-speakers), often underpinned by nationalist sentiments and stereotypes.

A statistical analysis performed using the Chi-square test reveals statistically significant relationships between attitudes and certain socio-biographical and linguistic variables. Specifically, the association between age and attitudes was statistically significant ($\chi2 = 11.21$, $p = 0.024$, $df = 4$) suggesting that participants over the age of 50 are more likely to express negative views as opposed to the younger generations who are much more likely to hold positive views (see Table 10). Participant sex, on the other hand, was not found to be statistically significant,

Table 10 about here

Table 11 about here

A statistically significant relationship also exists between the attitudes and the ethnic background of the participants ($\chi^2 = 6.054$, $p = 0.048$, $df = 2$) with white British more likely to produce negative and mixed responses compared to participants from other ethnic backgrounds (see Table 11). The category 'Other' includes participants from the following ethnic groups: White Other, Mixed, Indian, Pakistani, Other Asian, Black African, Chinese and Other.

Table 12 about here

The association between the length of time living in the UK and attitudes is also statistically significant ($\chi^2 = 8.397$, $p = 0.015$, $df = 2$). Those who had lived in the UK for their whole life reported less positive opinions (55.3%) than those who had spent only a part of their life in the

country (73.4%) (see Table 12). The former also have a tendency to report more negative and mixed attitudes.

The strongest association could be observed between attitudes and being multilingual vs. monolingual ($\chi^2 = 16.54$, p = 0.00, df = 2). Multilingual participants are much more likely to hold positive views towards living in a multilingual speech community (77.2%), compared to monolinguals (52.7%). In addition, the latter tend to report more negative and mixed opinions than multilinguals (see Table 13).

Linking mediatised representations with public views

Table 14 below summarises the participants' evaluation of 8 statements based on the dominant media discourses identified in the corpus study (in most cases with the actual press wording). The results are grouped into the overarching categories. The first category includes views in relation to *Multilingualism-as-an-opportunity Discourse* (responses to statements 1, 2, 4, 5) and the second attitudes towards *Multilingualism-as-a-problem Discourse* (responses to statements 3, 6, 7, 8).

Multilingualism-as-an-opportunity Discourse

The corpus analysis demonstrated that multilingualism is linked with employment, though this theme is less salient in recent years. When comparing this discourse with the views of our participants the majority seem to agree that multilingualism is beneficial, with 65% of the

respondents agreeing with statement 1 'Multilingual people can get better jobs'. The benefits of multilingualism are also expressed in statement 5 'Multilingual children have better opportunities in later life'; 52% agreed with this.

Another dominant theme identified in recent press discourse was the potential of multilingualism to delay Alzheimer's. Statements 2 and 4 were used to assess the public's views towards this theme. 71% of the participants agreed that speaking more than one language helps to keep the brain healthy. Interestingly, when the participants were directly asked whether multilingualism can prevent conditions such as dementia, the majority, namely 55%, indicated uncertainty. However 30% seemed to agree, possibly showing some familiarity with the discourse of 'multilingualism is a cure against Alzheimer's'.

Multilingualism-as-a-problem Discourse

When it comes to formal education the corpus study revealed a problematic view of multilingualism in relation to multilingual immigrant children, and perceived 'problems' with literacy. The answers to statement 6 and statement 8 do not reflect this concern. 64% disagree with statement 6, and 72% disagree with statement 8. In addition, 67% of the participants also disagreed with statement 3 ('Multilingual children achieve low grades at school'), suggesting that the majority does not perceive multilingualism as an obstacle to academic progress.

Since formal education and foreign language teaching were foregrounded in the corpus analysis, statement 7 was used to explore attitudes towards the notion of 'elite-multilingualism'. The findings seem initially to disprove the ideology, as 49% of the participants disagreed. However, it is worth noticing that nearly one third of our respondents agreed with this statement and 22% were uncertain.

Who is multilingual?

We deemed it necessary to investigate the participants' understanding of who can be considered multilingual, specifically in relation to elite multilingualism, as this was a pervasive theme in the press discourse. Five profiles of multilingual speakers were therefore designed using languages and qualifications that were identified as keywords in the corpus study. We are aware that the profiles might seem to reinforce certain stereotypes, but creating scenarios that were not consonant with the media representations would miss the point of this study. The profiles and the participants' answers are presented below.

Profile 1: Anna has just recently moved to Reading from Poland. She works as a cleaner in the local hospital. She speaks enough English to perform her daily duties but is not able to write in it.

Figure 3 about here

Profile 2: Leonie is a native French speaker who has recently moved to Reading from Paris to take up a position as a French teacher in a local secondary school. Her written English is good as she has an equivalent of A-level in English. However, she often struggles to understand and to communicate verbally in English with her colleagues and students.

Figure 4 about here

Profile 3: Ahmat was born in Reading to parents who came originally from Pakistan. He grew up in an Urdu-speaking household. Urdu, in which he has only a speaking ability, is the language he uses to communicate with his family. English is the language which he uses outside home, e.g. in the school.

Figure 5 about here

Profile 4: Jessica was born in Reading and her first language is English. She has an A-level in German but rarely uses the language (only once or twice a year when she visits her British friends in Berlin). She can get by in German in everyday situations (in shops, restaurants) and have basic/small talk conversations with German friends.

Figure 6 about here

Profile 5: Thomas was born in Reading and his first language is English. He has just finished his degree in English and French at Reading University and will be moving to France to take up a job as teacher of English in a primary school.

Figure 7 about here

Anna in profile (1) is a native speaker of Polish and has a non-prestigious job. She is able to speak basic English but is not able to write in it. Only 24% of the participants considered Anna to be multilingual and 34% clearly indicated that she is not. On the other hand, Leonie in profile (2) is a French speaker with a formal qualification in English. She is a teacher and her occupation is therefore more prestigious than Anna's. More participants considered Leonie multilingual (namely 37%) compared to Anna, although both lack skills in either spoken or written language. This suggests that knowing 'elite' languages and having a more prestigious job together with a recognised qualification are features that 'people on the street' are more likely to associate with being multilingual.

The case of Ahmat in profile (3) is relatively clear. He learned both his languages at a very young age and uses them both daily. Ahmat was the person that most participants considered to be multilingual, with 86% choosing 'yes'.

Jessica's case is similar to Leonie's in that she has a qualification and uses two 'elite' languages. Compared to Leonie, Jessica can hold a basic conversation in her second language,

whereas Leonie cannot. Jessica's case is also similar to Anna's in that they can both have a basic conversation in their second language. Nevertheless, Jessica is considered multilingual by more participants than Anna (namely 40% whereas in Anna's case only 24%). Perhaps the fact that Jessica has a qualification in an 'elite' language has influenced the participants to think that she is multilingual.

Finally, Thomas in profile (5) is regarded by 71% of the participants as multilingual, which is perhaps influenced by the fact that he has a university degree in his 'elite' languages. Possibly an assumption is made between a university degree in a language and having multilingual status.

With the exception of Ahmat, all profiles involved people who would either speak or write, or who had different kinds of qualifications and jobs. Interestingly, but perhaps not surprisingly, those who can speak elite languages and have a formally recognised qualification in a language and a more prestigious job are more likely to be considered multilingual. These findings suggest that elite languages, schooling and qualifications are indeed linked with multilingualism whereas 'lower' status languages are not, thus pointing to the pervasiveness of the 'elite' multilingualism identified in the media representations.

CONCLUSIONS

The aim of this study was to showcase how an in-depth analysis of the metalanguage of *multilingualism* in both the media and public responses to a survey can offer a more comprehensive understanding of discourses and ideologies surrounding multilingualism and their public uptake. Studying the metalanguage of multilingualism in two public contexts, media and public views, can help us not only understand better the mediatisation of multilingualism and its wider effects, but also address the partiality of research based on one context only. It can thus guard against over- or underinterpretations. Our findings indicate that

some discourses evident in British newspapers are shared by the general public whereas others are not. For example, the media discourse associating multilingual immigrant children with 'problems' seems to be refuted by the general public. Overall, our results reveal that the majority of the participants have positive attitudes towards living in a multilingual community, despite some of the negative representations projected in the British media. Concerns were however expressed by some respondents (mostly monolingual, older and from a white British ethnic background who have lived in the UK their whole life), and these were linked with immigration and underpinned by nationalist sentiments. Our results confirm the pervasiveness of the ideology of elite multilingualism for the general public (see, for example, Heller 2007; Horner 2011) in that multilingualism seems to be consistently associated with elite language, formal qualifications and prestigious employment.

The methodological strength of this study is that it successfully triangulated two research methods, a corpus-assisted discourse study and a sociolinguistic survey, to offer a more comprehensive and rigorous picture of public discourses of multilingualism. Corpus linguists concerned with aspects of discourse tend to study textual media data only and rarely engage with the community at large. Our study shows how corpus linguistic methods can effectively be brought together with other linguistic approaches to provide much more nuanced insights into public discourses by including voices from 'people from the street'.

However, there are several caveats that need to be highlighted. First and foremost, this study is exploratory in nature. Although we found statistically significant associations between attitudes towards multilingualism and some socio-biographical variables including age, ethnicity, length of residency in the UK, and self-identification as monolingual or multilingual, these aspects need to be further investigated using a larger stratified sample of participants to ensure better validity and generalisability of findings. We are also aware of the limitations of surveys in data collection. It is possible that in some cases, for example for reasons of political

correctness, respondents might have reported more socially desirable attitudes as opposed to what they truly believe, and may have hidden negative opinions. An online survey could reduce the effects of social desirability in self-reporting, (Wilson & Dewaele 2010) but it requires careful targeting of participants.

On a final note, we are consciously optimistic that despite the current negative climate surrounding immigration in post-Brexit Britain so often reinforced by some of the British media, the public is generally positive towards multilingualism. Nevertheless, much more work still needs to be done in order to promote multilingualism and counteract some of the negative associations. Our results could be of relevance to the general public, teachers and campaigners who work in the context of bi- or multilingualism. A successful campaign should emphasise the benefits that people link with multilingualism, such as better employment opportunities or cognitive advantages. It should also challenge some of the myths that are, at times, associated with multilingualism, especially regarding the linguistic and academic abilities of immigrant children. We hope that our study will offer a useful contribution to this work and debate.

REFERENCES

Agha, Asif (2011). Meet mediatization. *Language & Communication* 31 (3): 163-170.

Androutsopoulos, Jannis (2010). Ideologizing ethnolectal German. In Sally Johnson & Tommaso Milani (eds.), *Language ideologies and media discourses*, 182-204. London: Continuum.

Androutsopoulos, Jannis (2014). Mediatization and sociolinguistic change. Key concepts, research traditions, open issues. In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 3-48. Berlin, Boston: de Gruyter.

Baker, Colin (2008). Survey methods in researching language and education. In Kendall A. King (ed.), *Encyclopedia of language and education,* 55-68. New York: Springer Science and Business Media LLC.

Baker, Paul, Gabrielatos, Costas & McEnery, Tony (2013). *Discourse analysis and media attitudes.* Cambridge: Cambridge University Press.

Balam, Osmer & de Prada Pérez, Ana (2017). Attitudes toward Spanish and code-switching in Belize: Stigmatization and innovation in the Spanish classroom. *Journal of Language, Identity & Education* 16 (1): 17-31.

Blackledge, Adrian (2004). Constructions of identity in political discourse in multilingual Britian. In Aneta Pavlenko & Adrian Blackledge (eds.), *Negotiations of identity in multilingual contexts*, 68-92. Clevedon, UK: Multilingual Matters.

Blagden, David (2017). Britain and the world after Brexit. *International Politics* 54 (1): 1-25.

Blommaert, Jan (2005). *Discourse: A critical introduction*. Cambridge: Cambridge University Press.

Blommaert, Jan, Kelly-Holmes, Helen, Lane, Pia, Leppänen, Sirpa, Moriarty, Máiréad, Pietikäinen, Sari & Piirainen-Marsh, Arja (2009). Media, multilingualism and language policing: An introduction. *Language Policy* 8 (3): 203-207.

Bourdieu, Pierre (1991). *Language and symbolic power*. Harvard: Harvard University Press.

Caruana, Sandro & Lasagabaster, David (2011). Using a holistic approach to explore language attitudes in two multilingual contexts: The Basque Country and Malta. In Cecilia Varcasia (ed.) *Becoming multilingual: Language learning and language policy between attitudes and identities,* 41-67. Bern: Peter Lang.

Demmen, Jane, Semino, Elena, Demjen, Zsófia, Koller, Veronika, Hardie, Andrew, Rayson, Paul & Payne, Sheila. (2015). A computer-assisted study of the use of violence metaphors

for cancer and end of life by patients, family carers and health professionals. *International Journal of Corpus Linguistics*, 20 (2): 205-231.

Dewaele, Jean-Marc & McCloskey, James (2015). Attitudes towards foreign accents among adult multilingual language users. *Journal of Multilingual and Multicultural Development* 36 (3): 221-238

Dewaele, Jean-Marc & Wei, Li (2014). Attitudes towards code-switching among adult mono- and multilingual language users. *Journal of Multilingual and Multicultural Development* 35 (3): 235-251.

Duchêne, Alexander & Heller, Monica (2011). *Pride and profit: Language in late capitalism*. Oxford & U.K.: Routledge.

Fairclough, Norman (1989). *Language and Power*. London: Longman.

Gabrielatos, Costas & Baker, Paul (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996-2005. *Journal of English Linguistics* 36: 5-38.

Gardner-Chloros, Penelope, McEntee-Atalianis, Lisa & Finnis, Katerina. (2005). Language attitudes and use in a transplanted setting: Greek Cypriots in London. *International Journal of Multilingualism* 2 (1): 52-80.

Garrett, Peter (2010). *Attitudes to language: Key topics in sociolinguistics*. Cambridge: Cambridge University Press.

Giles, Howard (1970). Evaluative reactions to accents. *Educational Review* 22 (3): 211-227.

Heller, Monica (1999). *Linguistic minorities and modernity: A sociolinguistic ethnography*. New York: Longman.

Heller, Monica (2003). Globalization, the new economy, and the commodification of language and identity. *Journal of Sociolinguistics* 7 (4): 473-92.

Heller, Monica (2007). The future of bilingualism. In Monica Heller (ed.), *Bilingualism: A Social Approach*, 314-345. London: Palgrave.

Hilmarsson-Dunn, Amanda & Mitchell, Rodamond (2011). Multilingual migrants in England: Factors affecting their language use. In Cecilia Varcasia (ed.), *Becoming multilingual: Language learning and language policy between attitudes and identities*, 65-86. Bern: Peter Lang.

Hoare, Rachel (2000). Linguistic competence and regional identity in Brittany: Attitudes and perceptions of identity. *Journal of Multilingual and Multicultural Development* 21 (4): 324-346.

Horner, Kristine (2011). Media representations of multilingual Luxembourg: Constructing language as a problem, resource, right or duty. *Journal of Language and Politics* 10 (4): 491-510.

Ibarraran, Amaia, Lasagabaster, Peter, & Sierra, Juan Manuel (2008). Multilingualism and language attitudes: Local versus immigrant students' perceptions. *Language awareness* 17 (4): 326-341.

Irvine, Judith & Gal, Susan (2000). Language ideology and linguistic differentiation. In Paul V. Kroskrity (ed.), *Regimes of language: Ideologies, polities, and identities*, 35-83. Santa Fe: School of American Research Press.

Jaffe, Alexandra (2007). Minority language movements. In Monica Heller (ed.), *Bilingualism: A Social Approach*, 50-70. London: Palgrave.

Jaffe, Alexandra (2011). Sociolinguistic diversity in mainstream media: Authenticity, authority and processes of mediation and mediatization. *Journal of Language and Politics* 10 (4): 95-119.

Jaworski, Adam, Coupland, Nicolas & Galasiński, Dariusz (2004). *Metalanguage: Social and ideological perspectives*. Berlin: Mouton de Gruyter.

Johnson, Sally & Ensslin, Astrid (2007). *Language in the media: Representations, identities, ideologies*. London: Bloomsbury.

Jones, Rodney (2013). *Health and risk communication*. London: Routledge.

Kelly-Holmes, Helen (2011). Sex, lies and thematising Irish: New media, old discourses? *Journal of Language and Politics* 10 (4): 43-66.

Kelly-Holmes, Helen & Milani, Tommaso (2011). Thematising multilingualism in the media. *Journal of Language and Politics* 10 (4): 467-489.

Kelly-Holmes, Helen & Pietikäinen, Sari (2014). Commodifying Sámi culture in an indigenous tourism site. *Journal of Sociolinguistics* 18 (4): 518-538.

Kerswill, Paul (2014). The objectification of 'Jafaican'. The discoursal embedding of Multicultural London English in the British media. In Jannis Androutsopoulos (ed.), *Mediatization and sociolinguistic change*, 427-456. Berlin & Boston: de Gruyter.

Kilgarriff, Adam (2005). Language is never ever ever random. *Corpus Linguistics and Linguistic Theory* 1 (2): 263-276.

Lasagabaster, David & Huguet, Ángel (2007). *Multilingualism in European bilingual contexts: Language use and attitudes.* Clevedon: Multilingual Matters.

Martin-Jones, Marilyn, Blackledge, Adrian & Creese, Angela (2012). Introduction: A sociolinguistics of multilingualism for our times. In Marilyn Martin-Jones, Adrian Blackledge & Angela Creese (eds.), *The Routledge handbook of multilingualism*, 1-26. Abingdon: Routledge.

Milani, Tommaso, Davies, Bethan & Turner, Will (2011). Unity in disunity: Centrifugal and centripetal forces of nationalism on the BBC Voices website. *Journal of Language and Politics* 10 (4): 587-613.

Novak-Lukanovič, Sonja & Limon, David (2014). Attitudes to bilingual education in Slovenia. *Current Issues in Language Planning* 15(4): 426-442.

Office for National Statistics (2011) *Census of Population*. Crown Copyright.

Partington, Alan, Duguid, Alison & Taylor, Charlotte (2013). *Patterns and meanings in discourse. Theory and practice in Corpus-Assisted Discourse Studies (CADS).* Amsterdam & Philadelphia: Benjamins.

Preston, Dennis (2004). Folk metalanguage. In Adam Jaworski, Nick Coupland & Dariusz Galasiński (eds.), *Metalanguage: Social and ideological perspectives*, 75-101. Berlin: de Gruyter.

Pritchard, Rosalind (2004). Protestants and the Irish language: Historical heritage and current attitudes in Northern Ireland. *Journal of Multilingual and Multicultural Development* 25 (1): 62-82.

Rychlý, Pavel (2008). A lexicographer-friendly association score. In Petr Sojka & Aleš Horák (eds.), *Proceedings of recent advances in Slavonic natural language processing*, RASLAN 2008, 6–9. Brno: Masaryk University.

Scott, Mike (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In Marina Bondi & Mike Scott (eds), *Keyness in texts*, 43-58. Amsterdam: Benjamins.

Stavans, Anat & Hoffman, Charlotte (2015). *Multilingualism: Key topics in sociolinguistics*. Cambridge: Cambridge University Press.

Weisser, Michael (2016). *Practical corpus linguistics: An introduction to corpus-based language analysis*. Oxford: Blackwell.

Wilson, Rosemary & Dewaele, Jean-Marc (2010). The use of web questionnaires in second language acquisition and bilingualism research. *Second Language Research* 26: 103-123.

Vessey, Rachelle (2016). *Language and Canadian media: Representations, ideologies, policies*. London: Palgrave.

Appendix 1: Press sources and the number of articles obtained from each decade

Source	MinD1 1990-1999	MinD2 2000-2009	MinD3 2010-2013	Total
Daily Mail and Mail on Sunday	9	34	10	53
The Daily Telegraph	0	23	31	54
The Express	0	9	6	15
The Guardian	45	61	21	127
The Independent	72	41	20	133
The Mirror/The Mirror on Sunday	3	11	8	22
The Observer	9	14	14	37
The Sun	0	11	8	19
The Sunday Express	0	5	3	8
The Sunday Telegraph	0	11	5	16
The Sunday Times	16	29	17	62
The Times	41	64	21	126
Total	195	313	164	672

TABLES and FIGURES

TABLES:

Table 1: Corpus Data

Corpus	Words	Articles
MinD1 (1990 – 1999)	204,677	195
MinD2 (2000 – 2009)	437,006	313
MinD3 (2010 – 2014)	209,841	164
Total	851,524	672

Table 2: Participants

Age groups	16-24 (72) 37%		25-49 (97) 50%		50+ (24) 13%
Gender	Female (137) 70%			Male (57) 30%	
Ethnicity	White British (122) 63%			Other (72) 37%	
Years living in the UK	0-1 (6) 3%	2-5 (50) 26%	6-10 (17) 8%	11+ (39) 19%	Whole life (85) 44%
Monolingual or multilingual	Monolingual (93) 48%			Multilingual (101) 52%	

Native language	English (129) 66%		Other (65) 34%	
Number of languages known	One (93) 48%	Two (44) 23%	Three or more (57) 29%	

Table 3: Keywords in the MinD subcorpora

MinD1 (1990-1999)		MinD2 (2000-2009)		MinD3 (2010-2014)	
Keyword	Keyness Score	Keyword	Keyness Score	Keyword	Keyness Score
BILINGUAL	20.2	LANGUAGE	21.2	LANGUAGE	26.8
LANGUAGE	16.7	BILINGUAL	19.3	BILINGUAL	25.8
LANGUAGES	15.3	LANGUAGES	18.4	LANGUAGES	20.4
ENGLISH	10.2	ENGLISH	10.7	ENGLISH	11.7
FRENCH	9.6	FRENCH	10.7	FRENCH	9.7
WELSH	7.2	WELSH	9.3	SCHOOL	8.6
SPEAK	7.2	GAELIC	9.2	SPEAKING	8.1
SCHOOLS	6.4	SPEAK	7.6	CHILDREN	8
BILINGUALISM	6.1	CHILDREN	6.1	SPEAK	7.2
EDUCATION	5.8	SAYS	5.9	SCHOOLS	7
MULTILINGUAL	5.6	SCHOOL	5.3	PUPILS	6.7
QUEBEC	5.3	SPEAKING	5.2	BILINGUALISM	6.4
CHILDREN	5.2	SPANISH	5	SPANISH	6.4
TEACHERS	4.9	FOREIGN	5	FOREIGN	6.3
FOREIGN	4.9	MULTILINGUAL	5	ALZHEIMER	6.0

Table 4: Keywords in MinD1 grouped into semantic categories

Semantic Category	MinD1 Keywords (1990 -1999)
Education	schools, education, school, learn, learning, teaching, skills, taught, curriculum
Countries/regions	Quebec, Canada, Wales, France, Britain, Canada's, European, California
Social actors	children, teachers, pupils, students, Canadians, bilinguals, parents, linguists
Languages	English, French, Welsh, Gaelic, Spanish, German, Italian, Spanglish, Russian
Language-related terms	bilingual, language, languages, bilingualism, multilingual, grammar, trilingual
Descriptor/evaluation	foreign, fluent, ethnic, fluently, cultural, anti, multi, federal, official, poor
Employment	secretarial, jobs, secretary, salaries, recruitment
Communication	speak, speaking, says, translation, reading
Cities	London, Amsterdam, Birmingham, Paris
Medical/bodily terms	dyslexia, dyslexic, deaf
Others	year, sign, signs, demand, embassy, province, most, agency, society

Table 5: Keywords in MinD2 grouped into semantic categories

Semantic Category	MinD2 Keywords (2000-2009)
Education	school, learning, schools, learn, primary, education, teaching, taught, lessons,
Social actors	children, pupils, speakers, people, graduates, parents, teachers, immigrants

Languages	English, French, Welsh, Gaelic, Spanish, Catalan, Irish, German, Polish, Italian
Countries/regions	EU, Wales, UK, France, Belgium, China, Gaeltacht, Britain, Spain, Ireland
Language-related terms	language, bilingual, languages, multilingual, bilingualism, linguistic, minority
Descriptor/Evaluation	foreign, fluent, native, cultural, global, ethnic, national, multicultural, official
Communication	speak, says, speaking, translation, spoken, communicate, translated
Cities	London, Beijing, Brussels, Paris
Employment	job, business, career, employers
Medical/bodily terms	deaf, age
Others	signs, globalisation, culture, friends, international, immigration, tourism, Olympics

Table 6: Keywords in MinD3 grouped into semantic categories

Semantic Category	MinD3 Keywords (2010 – 2014)
Education	school, schools, learning, primary, education, immersion, teaching, lessons, taught
Social actors	children, pupils, students, speakers, teachers, professor, immigrants, Bialystok, Matras
Languages	English, French, Spanish, Polish, Urdu, Mandarin, German, Flemish, Chinese, Italian
Medical/bodily terms	Alzheimer, dementia, brain, cognitive, neurogenesis, ageing, age, onset, disease
Language-related terms	language, bilingual, languages, bilingualism, multilingual, linguistic, monolingual
Countries/regions	EU, Malta, UK, France, Belgium, Flanders, England, Poland
Descriptor/Evaluation	foreign, fluent, immune, cognitive, better, cultural
Communication	speaking, speak, says, spoken, speaks, translation
Cities	Brussels, Manchester, London, Peterborough, Paris, Toronto
Others	benefits, culture, cultures, Ofsted, ability, immigration, years

Table 7: The 10 strongest collocations of 'bilingual' in MinD3

Collocate	Freq.	LogDice	Collocate	Freq.	LogDice
education	41	10.795	stream	19	10.068
school	62	10.750	found	20	9.840
children	58	10.721	first	22	9.768
primary	30	10.425	better	19	9.758
being	25	10.103	English	21	9.744

Table 8: The 10 strongest collocations of 'English' in MinD3

Collocate	Freq.	LogDice	Collocate	Freq.	LogDice
speak	145	11.748	and	206	10.346
language	185	11.134	use	39	10.346
first	50	10.678	learn	40	10.204
their	66	10.526	children	46	10.066
French	51	10.439	not	61	10.057

Table 9: The 10 strongest collocations of 'Alzheimer' in MinD3

Collocate	Freq.	LogDice	Collocate	Freq.	LogDice
disease	34	12.784	against	11	11.076
patients	11	11.320	drug	7	11.026
onset	9	11.193	stop	7	10.962
diagnosed	8	11.142	delaying	5	10.664
symptoms	8	11.105	speaking	4	10.272

Table 10: Attitudes in relation to age

| | | | Attitudes towards multilingualism | | | |
			Positive	Negative	Both	Total
Age	16-24	N	50	6	16	72
		%	**69.4%**	**8.3%**	**22.2%**	**100%**
	25-49	N	69	10	19	98
		%	**70.4%**	**10.2%**	**19.4%**	**100%**
	50+	N	8	3	12	23
		%	**34.8%**	**13.0%**	**52.2%**	**100%**
Total		N	127	19	47	193
		%	**65.8%**	**9.8%**	**24.4%**	**100%**

Table 11: Attitudes in relation to ethnicity

| | | | Attitudes towards multilingualism | | | |
			Positive	Negative	Both	Total
Ethnicity	White British	N	72	14	36	122
		%	**59.0%**	**11.5%**	**29.5%**	**100%**
	Other	N	55	5	12	72
		%	**76.4%**	**6.9%**	**16.7%**	**100%**
Total		N	127	19	48	194
		%	**65.5%**	**9.8%**	**24.7%**	**100%**

Table 12: Attitudes in relation to length of time living in the UK

| | | | Attitudes towards multilingualism | | | |
			Positive	Negative	Both	Total
Time living in UK	Not whole life	N	80	6	23	109
		%	**73.4%**	**5.5%**	**21.1%**	**100%**
	Whole life	N	47	13	25	85
		%	**55.3%**	**15.3%**	**29.4%**	**100%**
Total		N	127	19	48	194
		%	**65.5%**	**9.8%**	**24.7%**	**100%**

Table 13: Attitudes in relation to monolingual vs. multilingual

		Attitudes towards multilingualism			
		Positive	Negative	Both	Total
Multilingual	N	78	3	20	101
	%	**77.2%**	**3.0%**	**19.8%**	**100%**
Monolingual	N	49	16	28	93
	%	**52.7%**	**17.2%**	**30.1%**	**100%**
Total	N	127	19	48	194
	%	**65.5%**	**9.8%**	**24.7%**	**100%**

Table 14: Participants' understanding of multilingualism in relation to the main public

discourses evident in British newspapers

	Strongly Agree	**Agree**	**Neither Agree nor Disagree**	**Disagree**	**Strongly Disagree**
1) Multilingual people can get better jobs	**32 (17%)**	**95 (49%)**	45 (23%)	19 (10%)	3 (1%)
2) Using multiple languages helps to keep the brain healthy	**38 (20%)**	**99 (51%)**	46 (24%)	8 (4%)	3 (1%)
3) Multilingual children achieve low grades at school	1 (0.5%)	1 (0.5%)	62 (32%)	**89 (46%)**	**41 (21%)**
4) Multilingualism prevents conditions such as Alzheimer's	12 (6%)	47 (24%)	**106 (55%)**	25 (13%)	4 (2%)
5) Multilingual children have better opportunities in later life	**17 (9%)**	**83 (43%)**	69 (36%)	22 (11%)	3 (1%)
6) Multilingual children mix their languages when they speak, and this means that they are confused	3 (2%)	25 (13%)	41 (21%)	**90 (46%)**	**35 (18%)**
7) Multilingualism is only helpful if the languages used are widely spoken	2 (1%)	55 (28%)	43 (22%)	**68 (35%)**	**26 (14%)**
8) Multilingual children will develop more slowly because they have to master more than one language	0 (0%)	14 (7%)	40 (21%)	**108 (56%)**	**32 (16%)**

FIGURES

Figure 1: Negation in the vicinity of the collocation pair 'English' + 'first'

436,000 secondary school pupils in the UK <u>did not have</u>	**English**	as their *first* language. "I'd be surprised if there was
figures show 80.8% of those whose *first* language <u>is not</u>	**English**	gained at least five GCSEs at grade C or above last year
school in the country where <u>none of its children</u> speak	**English**	as their first language. This fact fascinates <u>and repels</u>
This lesson is already going over my head. <u>Not speaking</u>	**English**	as a first language <u>could amplify differences of ability</u>
School, Peterborough, where <u>not one pupil</u> speaks	**English**	as a first language," <u>thundered</u> Peter Hill in The Express
blighted prospects for children whose first language <u>is not</u>	**English**	<u>warned</u> Tory councillor Imtiaz Ameen. 'Of children are

Figure 2: *'Overall, would you say that living in a town like Reading where lots of languages are spoken is a positive or negative thing?'*
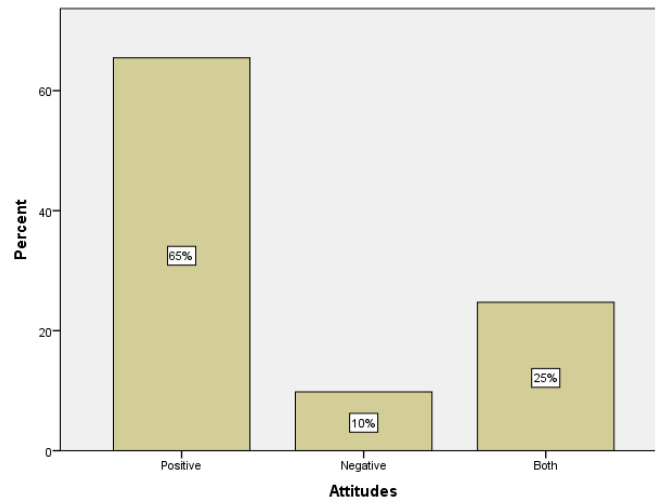
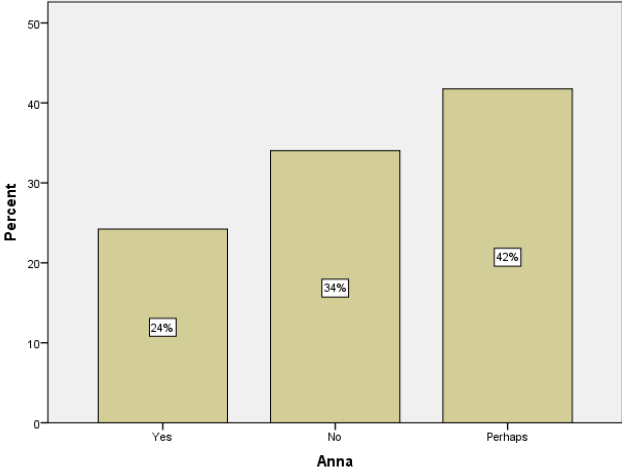Figure 3: Do you consider Anna to be multilingual?

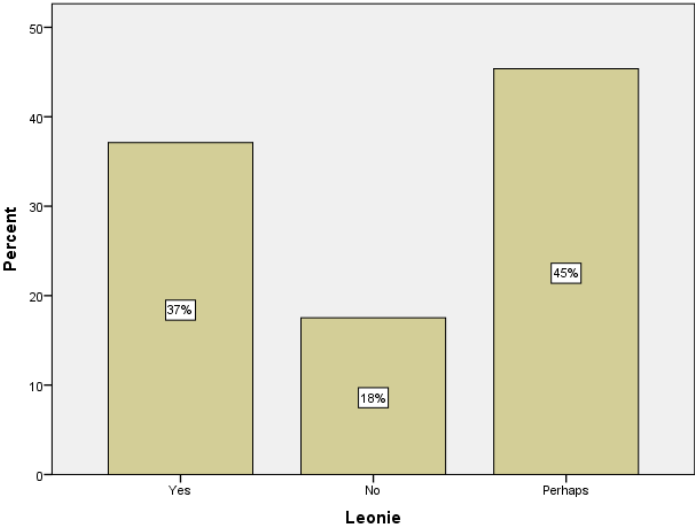Figure 4: Do you consider Leonie to be multilingual?
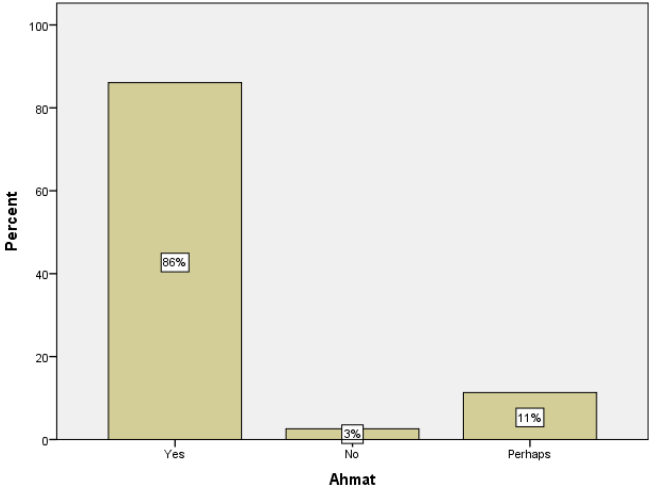
Figure 5: Do you consider Ahmat to be multilingual?

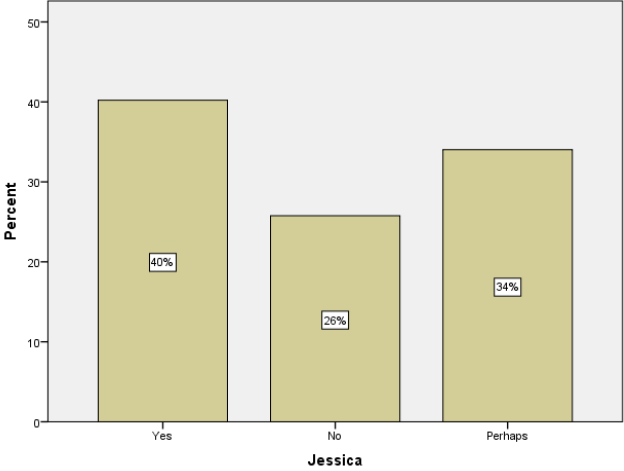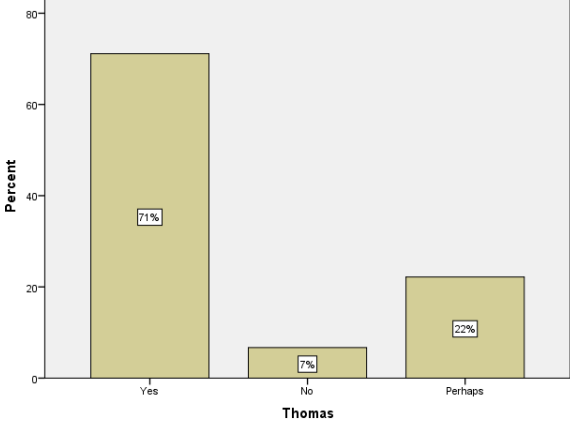Figure 6: Do you consider Jessica to be multilingual?

Figure 7: Do you consider Thomas to be multilingual?

[ii] Regular expressions (often referred to as regex) were created as part of the formal language theory in computer science and involve characters that can be combined into sequences or syntaxes to perform complex searches in large collections of texts. Weisser (2016) offers a useful introduction to regular expressions and their use in corpus linguistics.